Thesis/
Reports
Daniel,
T.C.

ACCEPTABILITY THRESHOLDS FOR PUBLIC LAND MANAGEMENT : SCIENTIFIC AND POLICY PERSPECTIVES

By Terry C. Daniel

College of Social and Behavioral Sciences
Department of Psychology

THE UNIVERSITY OF
ARIZONA®
TUCSON ARIZONA

P.O. Box 210068
Tucson, AZ 85721-0068

March 12, 2003

To: Dr. Tom Brown
USDA Forest Service
Rocky Mountain Research Station

Re: Final Report
#28-JV7-961

From: Terry C. Daniel
Professor of Psychology and
Renewable Natural Resources

cc: Carolyn J. Lumar
Lead Grants & Agreement Specialists
Rocky Mountain Research Station

Douglas Hilyard
Sponsored Projects Services
University of Arizona

Attached is the final report for our cooperative project, *Theoretical development of the threshold of acceptability*. I have titled the report, *Acceptability thresholds for public land management: scientific and policy perspectives*, consistent with the direction of our many discussions over the years of (and following) this project. As we are both aware, this paper hardly puts the issues to rest, but I believe that it does meet the formal requirements of our research agreement. I am hopeful that it will also provide a stimulus for continuing our consideration of the appropriate roles for "thresholds" and "acceptability" in environmental science and policy-making.

Acceptability thresholds for public land management: scientific and policy perspectives

Final Report for

Theoretical development of the threshold of acceptability.

Cooperative Research Agreement #28-JV7-961
with
Rocky Mountain Forest and Range Experiment Station

Submitted by

Terry C. Daniel
University of Arizona

Acceptability thresholds for public land management: scientific and policy perspectives

Terry C. Daniel
University of Arizona

Introduction

Few would quarrel with the idea that environmental management policies should be science based. Most would take that to mean that decisions about actions on the environment should be based on the best available biological, ecological and physical science. The principle role for science in the development of environmental management policy is to anticipate (project) the biological and physical consequences (outcomes) of alternative management actions. For example, bio-physical science can provide projections/estimates of the amounts of soil erosion, the rate of tree growth, and the numbers of wildlife species supported or displaced by any given policy. Through understanding of relevant processes, science may also suggest ways of altering those effects. A smaller number of people (including some policy makers) recognize that human/social science also has a role to play, especially when public lands/environments are involved. That is, if bio-physical science should be used to anticipate the effects of any contemplated action on the bio-physical world, then human/social science should be used to anticipate the effects of environmental management on people. For example social sciences can provide projections of changes in income amounts and distributions, numbers of jobs produced or eliminated, user-days of recreation activity supported or increases (decreases) in commuting time associated with alternative policies.

So long as projection of outcomes is focused on relatively concrete bio-physical or social indicators (tons of soil lost, numbers of jobs produced), traditional objective-positivist models of science fit fairly comfortably with the outcome "assessment" components of policy making. There is somewhat less comfort (and clarity) when the assessment of outcomes borders onto "evaluation," as when the outcome to be projected is the *quality* of habitat produced or the *quality* of recreation opportunity provided. Even here, relatively uncontroversial, scientific methods have been found to assess levels of quality within an outcome dimension. For example, reasonably objective/scientific assessments of relative habitat quality can be made for a restricted class of wildlife species, or for a limited set of recreation opportunities. But traditional bio-physical and social science methods and models are less suited for assessments requiring (or implying) valuations across classes of outcomes (such as determining habitat quality across competing wildlife species, comparing the value of wildlife habitat to jobs, or balancing recreation opportunities against damage to sensitive ecosystems).

Science of value

The role for science is less clear, and less generally accepted, when the task is to determine which of a number of potential multidimensional policy outcomes are "better" than others. Traditional bio-physical science, and even most social science has less to

offer as the policy task shifts from *assessment* of outcomes to the *evaluation* of outcomes. What is needed here is a *science of value*.

Among contenders for a science of value, economics is the most developed and rigorous, and the most often cited in formal government (and business) decision processes (even if it is almost equally often ignored or misused in the final setting of policy). Other contenders might include political science, where value is denominated in power and traded between institutions. There are also "community-based planning" advocates who argue for a "qualitative" case-by-case process of negotiation and consensus building among local stakeholders. By this approach, outcomes and their values are adamantly not measured, and success is indicated by agreement among the participants that the decision making process was fair and contributed to a healthier, happier, more harmonious, more sustainable community. The multi-discipline of decision science is another contender for a value science. This approach employs a hybrid mix of economic, psychological and system science principles in a rigorous and usually mathematical process to articulate individual (and sometimes group) "values" that underlie choices among policy options. Like economics, decision science usually takes the individual to be the source and arbiter of value—it is individual *preferences* that determine which option is "best." But decision science does not require that value be denominated in dollars (or even that it be uni-dimensional), and it more easily accommodates human failures to follow prescribed rules and formal logic of valuation.

Value and the acceptability threshold

Faced with the ambiguity of policy evaluation and the lack of a clear consensus on how it is to be accomplished--using science or some other method/model--it is little wonder that policy makers are somewhat desperate for a "defensible" way to determine the "best" among management options. In this context, there has been some attraction to the notion of a "threshold of acceptability." In the more extreme form this approach hopes to derive evaluations directly from science-based outcome assessments. That is, to find some "scientific" basis for dividing outcomes (tons of soil lost, user-days of recreation opportunity, etc) into mutually exclusive "acceptable" and "unacceptable" categories. A variant of this approach is the "inherent value" assumption, where some environmental outcomes are taken to be so important that they trump all other considerations (such as threats to endangered wildlife species). A less radical (more sophisticated) version of the threshold approach is to seek some public consensus in classifying outcomes into acceptable (to the public) and unacceptable categories. Part of the appeal of the threshold approach is the assumed greater stability of binary value classifications (people know what they like and what they don't like) compared to the presumed greater volatility of finer grained value distinctions (by how much is option A favored over option B?). Another attractive feature of thresholds of acceptability is that once the acceptable ("good") class is defined, managers can exercise considerable freedom in selecting policies from within that class. Thus, for example, once it has been determined what range of outcomes are "acceptable," the manager is free to create and implement management options within that range based on expert and administrative criteria, without worrying further about "evaluation" (especially by the public).

Whether based directly on bio-physical or social features of outcomes or on some measure of public acceptability, the threshold approach requires some important assumptions. For example, large amounts of soil loss might be assumed to be "bad" and smaller amounts "good" (a not too controversial assumption in most cases). A more difficult requirement is that relevant bio-physical (or human-social) outcome variables, or the public acceptability variable, must exhibit some clear discontinuity over the range of policy options considered, so that a "threshold" is apparent (left panel below).

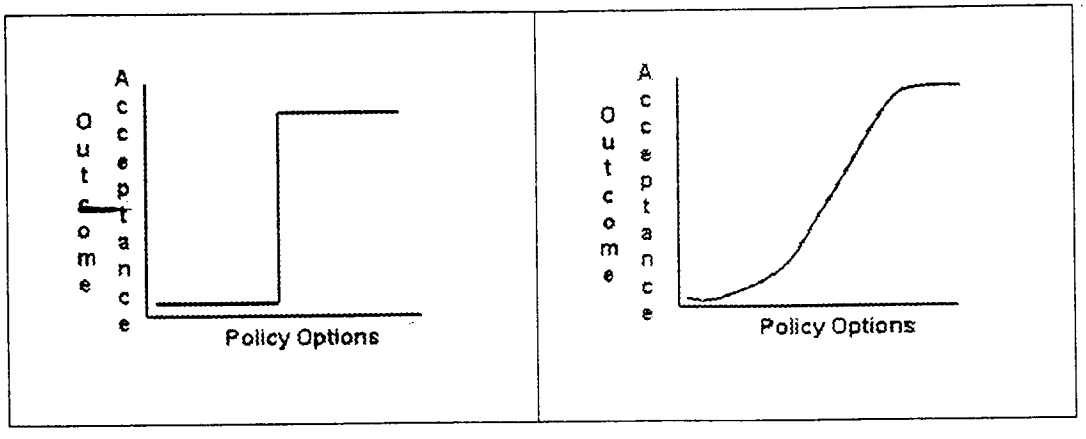

Figure 1. Outcome/acceptability functions

Unfortunately, both (bio-physical and social) outcome variables and public acceptance are almost always continuous over any reasonably wide and graded set of policy options (right panel). Moreover, outcomes and public acceptance functions tend to be relatively smooth (at least at finer levels of analysis), so that there are rarely any obvious discontinuities or inflection points that can unambiguously be claimed as the "acceptability threshold." As a consequence, outcome variables and public acceptance will almost always vary within the acceptable (and unacceptable) classes, contrary to the desire to treat them as homogeneous classes within which management discretion may be freely exercised.

Further discussion of the scientific and policy merits of acceptability threshold approaches to valuation are presented below in the context of a particular environmental management example. A simple experiment asked separate panels of untrained observers to judge the perceived *fire hazard* and the *scenic beauty* represented by hypothetical forest homesites, depicted by digital color photographs. A third set of observers rated the same sites for overall acceptability as a homesite, taking both scenic beauty and fire hazard into account. The study provides a simple model of a common environmental management/policy situation, where the acceptability of outcomes depends on tradeoffs among competing objectives. The analysis investigated the relationships among the two "outcome" dimensions (perceived fire hazard and scenic beauty) and judged acceptability, and explored and critically evaluated how several "threshold of acceptability" concepts might be applied in this context.

Evaluating forest homesites

Wildfire poses serious threats to homes and property in many forest areas. A national program of "fuel reduction" treatments (thinning and prescribed burning) seeks to reduce fire hazard in these areas, but some treatments can have adverse effects on scenic (and other) values that are also important to residents. Perception and evaluation of tradeoffs between fire hazard (safety) and scenic quality is central to public acceptance of fuel treatments near their homes. The study reported below explored public perception of fire hazard and scenic quality for hypothetical homesites in a fire prone forest environment.

Method

Sets of digital photographs were selected to represent views looking out from hypothetical forest homesites in a web-based perceptual survey. The sites were homogeneous with respect to topography (essentially flat) and did not exhibit any features such as lakes, streams or rock outcrops. Views were all "near-view" (within-stand). Vegetative fuel conditions depicted ranged from high to low fire hazard. Separate groups of observers independently rated each site for ***scenic beauty***, ***fire hazard*** or overall ***acceptability*** (as a forest homesite). Rating scale data were subjected to correlation-regression analyses to determine the separate contributions of aesthetic and hazard perceptions to overall homesite preferences. The suitability of several "threshold of acceptability" approaches was also evaluated.

Study Sites--Hypothetical study sites were created by selecting scenes (photographs) from a large set of southwestern ponderosa pine forest plots on the Coconino National Forest in northern Arizona. Plots were originally inventoried and photographed between 1979 and 1982. Each plot was centered around a sample point along a linear transect within a forest "stand," as delineated by professional Forest staff, and four color slides (90-degree separation) were taken from each plot center to represent visual features. Slides were subsequently commercially digitized and stored as compressed high-resolution JPEG files. The available forest scenes were inspected and a total of 160 scenes were selected to represent a wide range of apparent fuel load/fire hazard levels and scenic beauty values. These 160 scenes were organized into sets of 4 to create 40 homogeneous "artificial" forest homesites for the study. Thus, the 40 sites, 4 views per site, depicted hypothetical homesites that represented a wide range of fire hazard and scenic quality.

Study participants--Subjects were all undergraduate college students meeting research participation requirements for an introductory psychology class. Student populations have frequently been used in similar studies, and they were deemed appropriate for investigating basic aspects of hazard-amenity tradeoff processes. Experiment availability was advertised by a standard announcement posted among dozens of other experiments. A total of 79 subjects, participating in groups ranging from 3 to 10 (depending upon voluntary sign up rates) reported to a multi-station computer laboratory. Each subject was assigned to an individual PC workstation, where they independently and interactively viewed instructions and responded to the 40 hypothetical homesites. All subjects appearing for a given experimental session were assigned to the same rating scale (*scenic*

beauty, *fire hazard* or *acceptance*), so that any questions about procedures would not expose the other rating conditions in the experiment.

Perceptual ratings--A web-based survey procedure allowed multiple subjects to participate at the same time. Each subject was logged onto the appropriate form of the survey (rating condition) and each proceeded independently in a self-paced procedure to read instructions, observe preview images and respond to the 40 homesites. Each homesite was represented by four individual views/scenes. After each of the four views of a homesite had been self-presented at least once, the subject recorded his/her rating on the assigned scale for that homesite, based on all four views, and then proceeded to the next homesite. The order of the 40 homesites was individually randomized for each subject, and ratings were automatically ordered and entered into a relational database. After all 40 homesites were rated, a screen thanked the subject for his/her participation and a brief description of the purposes of the study was presented. Most subjects completed the rating procedure in approximately 30 minutes.

Rating scale conditions--The hypothetical homesites were all rated for *scenic beauty (SB)*, *fire hazard (FH)* and overall *acceptability (Acc)*. General instructions to all groups emphasized that advances in computer and communications technologies were supporting development of a dispersed workplace in which residential choices might be largely unconstrained by geography. In that context it was reported that increasing numbers of people were choosing to live in rural areas, especially forests, because of the scenic beauty, naturalness, privacy and other benefits afforded there. It was also noted that these same areas are subject to significant hazards, especially wildfire. In short, the general instructions established the fire-prone "wildland/urban interface" as the setting for the "homesites" to be evaluated.

Subjects were instructed to evaluate sites on the dimension (scale) to which they were assigned, with no reference to the other dimensions. *Scenic beauty* was defined in terms of "natural beauty" and "visual aesthetic quality" of the forest landscape. *Fire hazard* instructions included a brief description of how the volume/density and distribution ("ladders") of vegetative fuels affects wildfire behavior (intensity). Overall *acceptability* instructions emphasized the tradeoffs between aesthetic values (scenic beauty) and hazards (wildfire) in forest areas. Subjects were asked to evaluate the homesites based on their own assessment of the relative importance of these two factors, with all other factors (access, quality of home and utilities, proximity to lakes and other amenities) assumed equal.

Results

The first level of analysis focused on the relationships among the three scales. Simple correlations between scales (Figure 2) revealed expected relationships. *Acceptability* was strongly positively related to *scenic beauty* ($r = .88$) and less strongly negatively related to perceived *fire hazard* ($r = -.70$).

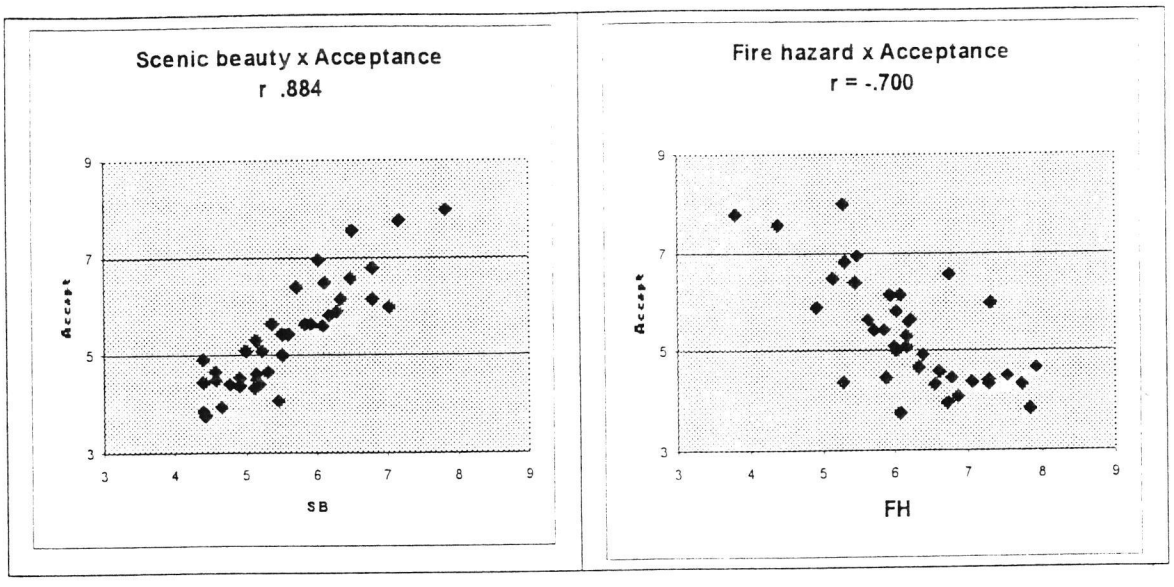

Figure 2. Correlations between ratings of *acceptability* and *scenic beauty* (left) and *fire hazard* (right).

The relationship between rated *scenic beauty* and *fire hazard* was also negative (r = -.55), but of only moderate strength. All three correlations are highly significant statistically, and the strength of relationships is substantial, especially considering that independent groups of subjects produced the three ratings.

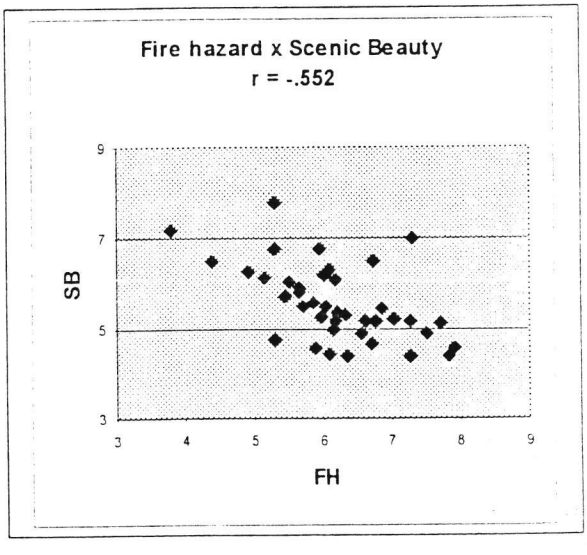

Figure 3. Correlation between ratings of *fire hazard* and *scenic beauty*.

Multiple regression analysis revealed that *acceptability* could be explained by the joint contributions of perceived *scenic beauty* and *fire hazard* ($F(2/ 37) = 101.99$, adjusted $R^2 = .84$). The standardized regression coefficients indicated that *scenic beauty* contributed substantially more to *acceptability* ($\beta = .72$) than did *fire hazard* ($\beta = -.30$).

Acceptability thresholds

The correlation-regression analysis above indicates that *acceptability* increases as scenic beauty increases and fire hazard decreases. This relationship is depicted in figure 4, below, where sites are ordered by mean acceptability rating.

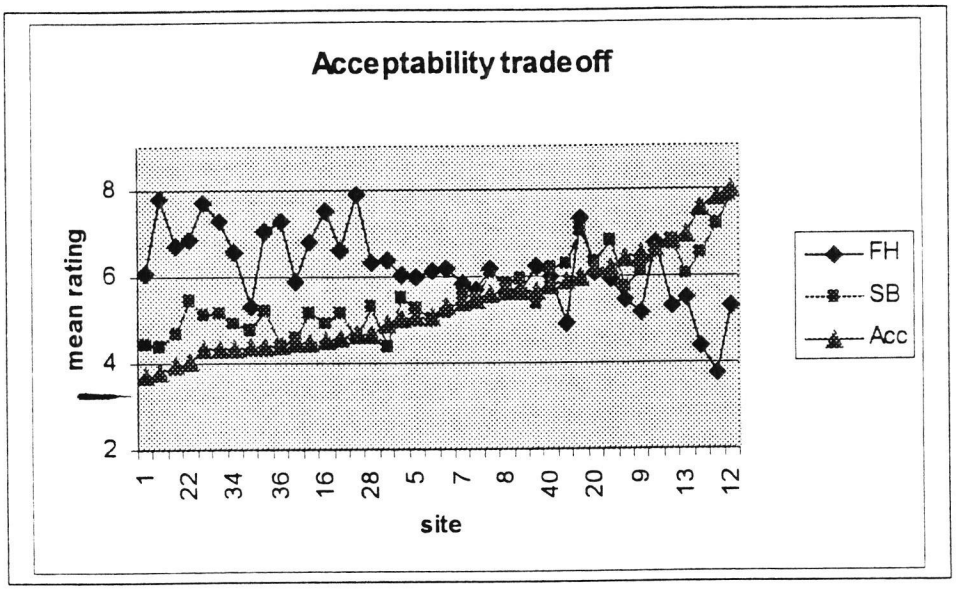

Figure 4. Acceptability trade off between perceived fire hazard and scenic beauty.

Inspection of figure 4 reveals that acceptability follows very closely the pattern of perceived scenic beauty, and less closely the inverse of fire hazard, as was precisely quantified in the multiple regression results reported above.

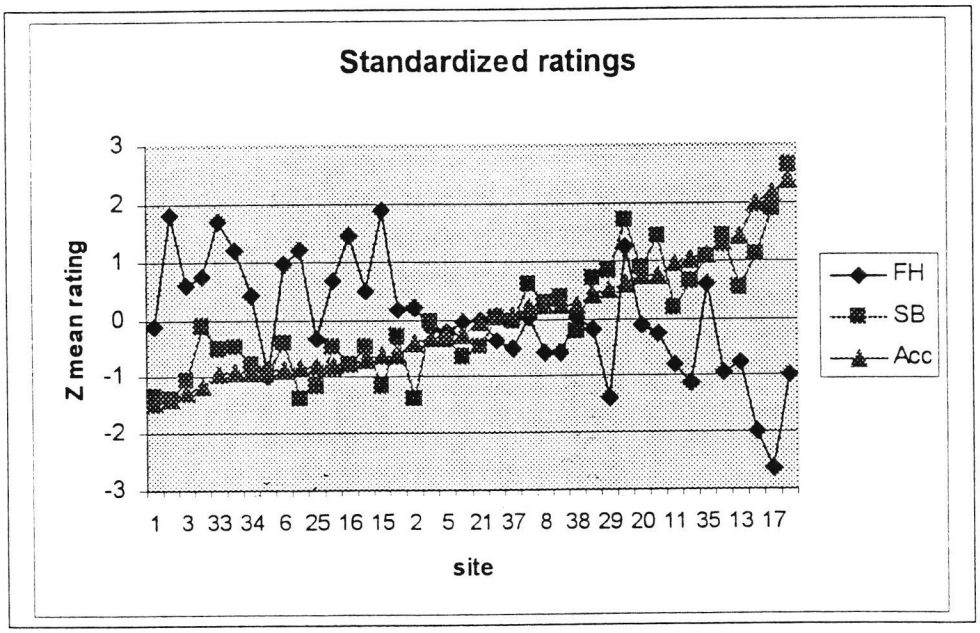

Figure 5. Standardized acceptability trade off between perceived fire hazard and scenic beauty.

In figure 5, the acceptability tradeoff is depicted in a "standard score" form. Mean ratings (per site) have been individually transformed to Z scores for each rating scale (based on the respective rating scale means and standard deviations), with sites again ordered by the standardized mean *acceptability* rating. The standardization procedure removes the arbitrary differences in scale origins ("zero") and intervals between the scales.

There is no indication in the data reported and illustrated above of any "inflection points" in acceptability, any points where perceived acceptability makes a "step" or "quantum change" from values below to values above. That is, inspection of the acceptability function over the range afforded by the sites selected for this study does not suggest any particular point that might be chosen as the "acceptability threshold" for forest homesites.

One might be tempted by the graphics to argue that the point at which the three scales intersect (at or near their respective means) represents some significant change in the tradeoff relationship—the point where scenic beauty and fire hazard are in some kind of balance. For these data, this point is approximately at the center of each of the three rating scales (5.5 on a 10 point scale), which might provide another rationale for establishing the threshold. The location of this intersection, however, is largely an artifact of the procedure for obtaining the measures.

The origin and interval properties of a rating scale are not "absolute" values with respect to the properties they are intended to measure (e.g., perceived *fire hazard, scenic beauty* or *acceptability*). The numbers used for a rating scale may be claimed to imply order (e.g., policies receiving a rating of 5 are more acceptable than policies rated a 4) and even indicate equal intervals across the range of the property being measured (e.g., the change in rating from 1 to 2 represents the same magnitude change in *acceptability* as that from a rating of 4 to 5), but the particular origin and interval characteristics of the scale are arbitrary with respect to the "scale" of the property being measured. For example, the investigator may choose to have observers rate options on a 5-point scale, a 10-point scale or a 100 point scale. None of these options can claim any special relationship to the underlying (psychological) dimension that is nominally being measured. Similarly, expressing a 7-point scale as "-3 to +3," (instead of 1 – 7) and identifying the middle value as "0" does not assure that the midpoint represents some absolute lack of the property measured, or some "neutral" or "absolute indifference" point that can be generalized beyond the particular context in which the measures are taken.

The set of options (forest sites/conditions) presented to observers in a rating scale procedure has substantial effects on the particular ratings assigned to each option. A given option might be rated higher or lower, depending upon the other options that are presented in the set being rated. Similarly, the amount of change in the measured property (e.g., the magnitude of difference in acceptability) required to produce a unit change in the overt rating response (e.g., from a rating of 4 to a rating of 5) will depend upon the range of the options offered. Thus, drawing conclusions about perceived properties of the environment based on ratings (or rankings or choices or magnitude

estimates) is complex, and is the subject of measurement theory and research in general, and psychophysics and psychometrics in particular.

Midpoints and means

"Thresholds" might be forced upon these continuous functions in a number of ways. For example, numerous investigators have chosen midpoints, means, or other arbitrary points on overt rating scales as representing some significant point on the underlying property scale. Figure 6 shows graphically how the nominal midpoint of the 10-point rating scales (5.5) might be designated as a "threshold" for acceptability.

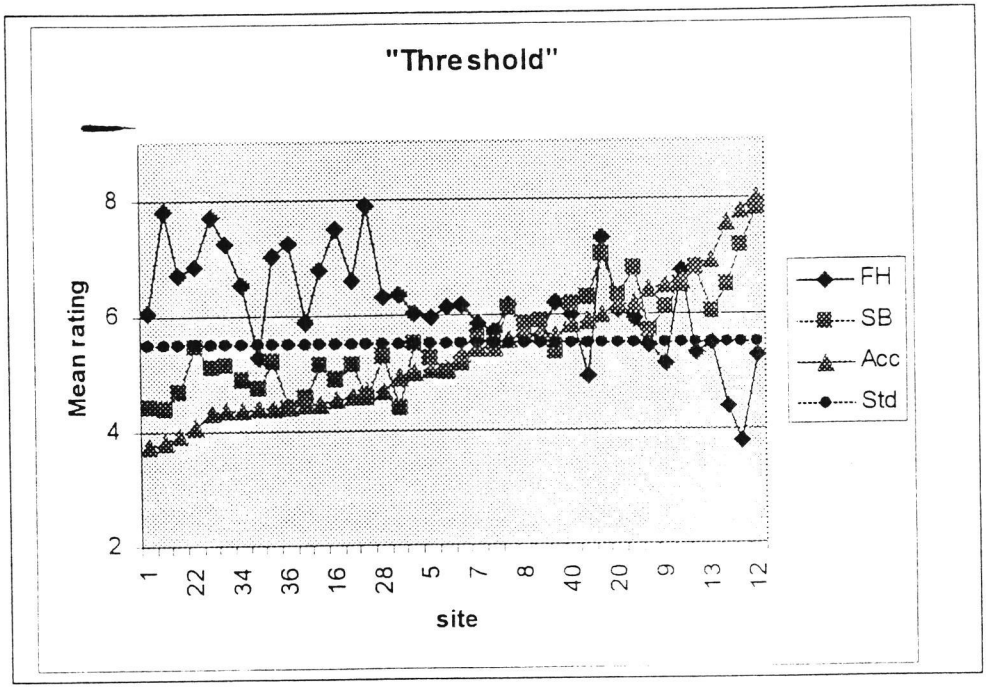

Figure 6. Representation of a "midpoint" criterion for an acceptability threshold.

The line across the graph at 5.5 on the Y axis (labeled Std) represents the nominal midpoint of the 10-point scales. By this approach, the "threshold for acceptability" (the "indifference" or "neutral" point) falls about half way between the conditions represented by hypothetical sites 37 and 39 (on the X axis, not all sites are listed), dividing the current option set into 23 "unacceptable" sites and 17 "acceptable" sites. As discussed above, while this point is a nominal "midpoint," it is quite arbitrary with respect to the underlying property represented (the "perceived acceptability" of forest management/outcome options). By this approach, the least of the "acceptable" options differs by less than 0.2 from the highest "unacceptable" option, but differs more than 10 times as much from the most acceptable option (in terms of the means of the nominal acceptability rating scale). Of course, one could expand the range of the "indifference" zone (by +/- one standard error, for example), but his would only create a similar arbitrary line between "neutral" and "acceptable" (or "unacceptable"), with no better scientific or practical justification.

Another approach might be to use the mean of acceptability ratings to define the "threshold." Figure 7 illustrates this approach by using the standardized ratings (so that the mean for all scales is zero).

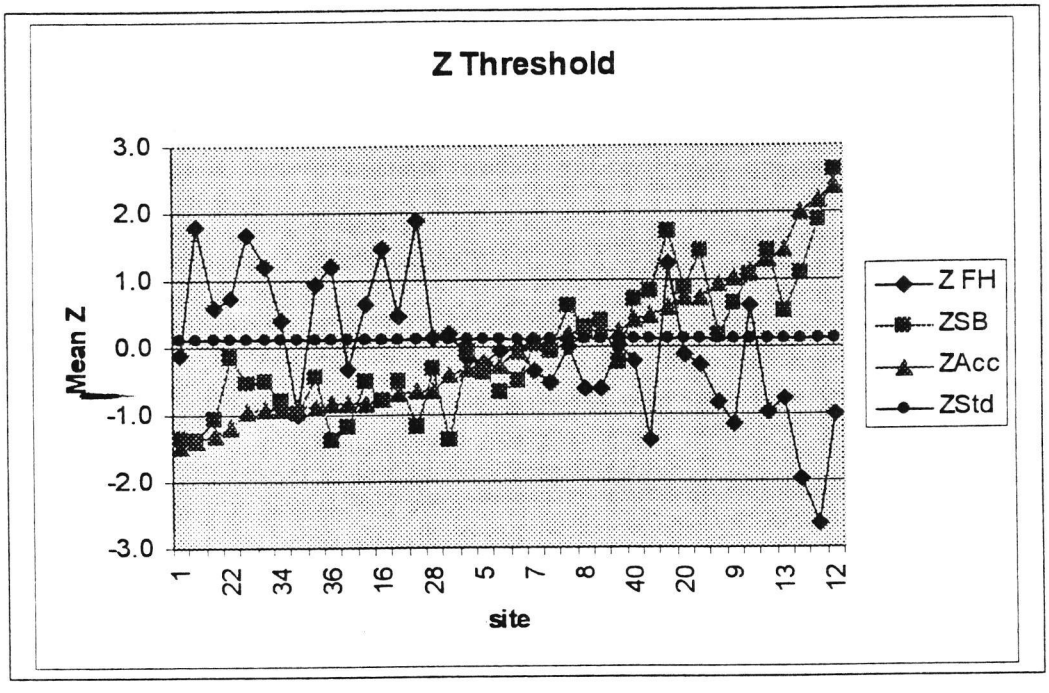

Figure 7. Representation of a "mean" criterion for the acceptability threshold.

By this criterion, the threshold falls between the conditions represented by sites 21 and 7 on the X axis (the same sites would define the threshold on the raw mean rating scales), resulting in 21 "unacceptable" and 19 "acceptable" sites. The threshold changes, relative to the midpoint criterion approach, but the issues of arbitrary origins and intervals, and the theoretical and practical ambiguity of any line (or lines) drawn on the continuous acceptability function are unchanged.

The midpoint and mean criterion approaches result in very similar "thresholds" (in terms of the forest condition/X axis), but the change does call attention to another disadvantage of a threshold approach. Note that there is no change in the acceptability function—that is the order and magnitude relationships indicated among the forest conditions is not affected by what criterion might be used to define the "threshold." This would be expected to hold in general—the relationship among options (the acceptability function) will be more stable than the location of any designated point along the function. It would be expected, for example, that adding some number of site options that offered "better" (more acceptable) conditions (and/or deleting some number of lesser options) to the set of alternatives rated could substantially shift the location of the "threshold" on the condition/X axis—whether the threshold is defined by the midpoint, the mean or some other "rational" point-criterion. That is, options that were rated near the midpoint (or mean) in the first set of options would be expected to obtain lower ratings in the altered set of options, so the "threshold of acceptability" would shift along the condition axis.

Options that were above the threshold in the first set ("acceptable") would be below the new threshold ("unacceptable"). At the same time, the relationships among the sites/options with respect to perceived acceptability would be largely unchanged—the acceptability function would be stable.

Policy implications

The advantage sought from an "acceptability threshold" for policy and planning purposes is a "clear line" between acceptable and unacceptable options, and the designation of a range of options (acceptable ones) within which managers are allowed to exercise their discretion (based on expert opinion and analysis), without further consideration of public preferences/values. This assumes that the distinction between acceptable and unacceptable is more robust than the continuous outcome/value functions on which any threshold must be based, and that, in contradiction to the usually continuous, monotonic underlying functions, the set of acceptable options is essentially homogenous with respect to relevant public values. The illustrations and discussion above suggests that contrary to this line of reasoning, thresholds will generally be less stable than the outcome/value functions on which they are based. Further, options within any set designated as acceptable (or unacceptable) by any threshold will typically differ in value, exactly according to the continuous value function. As a result, while acceptable options will always be higher in value than unacceptable options, there may often be more difference in value between two acceptable options than between an acceptable and an unacceptable option. Moreover, there is nothing in the threshold concept that absolves the manager of public lands from striving to maximize public value/acceptability, whether his/her choices are within or between acceptable and unacceptable options.

Epilogue

There is nothing in the above that argues against the use of set "standards" in environmental policy and management. It is generally clear in the setting of standards (as for air quality, water quality, etc) that the standard is based on processes and considerations separate from the assessment of the relevant outcome variables. For example, the standard for lead concentrations in drinking water is not determined by inspection of the distribution of lead concentrations, but from a complex political-legal negotiation process that takes many factors (health, economics, etc) into account. The standard is stated in terms of an allowed ("acceptable") lead concentrations, but no one argues that the level that is set represents natural "threshold" for lead in water, or that concentrations above that amount are homogeneously "unhealthy" and concentrations below it are "healthy."